

# 回答書

Dynamic anchorを用いた  
単眼画像での距離推定

60200053 佐々木寛史 6/13

# 質問への回答

- YOLOにおいて出力される特徴マップを表すSxSのグリッドの数を変更した際にどういった変化が起こるのか？

Sの値はデフォルトで13となっており、この値を増加させると予測されるBounding Box (BB)の数が多くなり、計算量が増えてしまうため処理が遅くなると考えられます。しかし一方で各BBの座標はセルごとに予測されるため、検出精度は高くなると考えられます。

- Dynamic anchorを出力する  $\sigma(d)$  の引数として角度を用いているが、その角度とはどのようなものか？

あくまで検討段階ですが、あくまでanchorを出力するための補助的な情報として与える方針です。（例：物体が斜め、横、縦向きのいずれかという情報）

- Dynamic anchorのメリットとデメリットは何か

膨大な数のanchorに関して距離の推定を行うことによって計算量が膨大になってしまうことを防ぐために、物体の特徴に応じたanchorを逐次用いることができるというメリットがある一方で、毎回適切なanchorを出力するとは限らないというデメリットがあると考えられます。

- 3D Bounding Boxのアノテーション付の負荷が大きいということだが、CGを用いることによってその課題を解決することは検討したか

現実世界のデータセットとCGのデータセットには画像の特徴に乖離があるため、CGによって学習を行い、現実世界で用いることは困難です。近年その乖離を緩和するための研究は盛んに行われていますが、CGのデータを作成すること自体にも負荷があるため、そういったデータを新たに用いることなく、既存法と同等の距離推定が可能になるように研究を行なっていきたいと思います。

- 車以外での距離推定は検討しているか。

現在では検討していません。歩行者など他の物体ではそれ特有の新たな問題点があるはずなので、余力があれば検討していきたいと思っています。

- Dynamic anchorのに先行研究はあるのか

すべての論文を確認したわけではないですが、CVPRなどの有名な学会で発表されている論文において類似した内容は見られませんでした。

- もう少し具体的に実験方法を知りたい

KITTIのベンチマークにおいて、推定された距離の絶対誤差平均を用いて従来法との比較を行います。その他にも、RecallやPrecisionにおいてステレオカメラを用いた手法などとの比較を行い、検出精度の評価も行います。なお、今回の質問において、動画を用いた検出結果のデモを用いることでこういった内容かというイメージがしやすいのではないかと考えたので、動画の作成にも取り組もうと思います。